NAVAL POSTGRADUATE SCHOOL

Monterey, California

D IONS

MARCH, 1992

Thesis Advisor: Murali Tummala

Approved for public release; distribution is unlimited

Approved for public release; distribution is unlimited.

Adaptive Filter Design
Using
Discrete Orthogonal Functions

by

Steven J. McPhillips
Lieutenant, United States Navy
B.S., United States Naval Academy, 1986

Submitted in partial fulfillment
of the requirements for the degree of

MASTER OF SCIENCE IN ELECTRICAL ENGINEERING

from the

NAVAL POSTGRADUATE SCHOOL
March 1992

ABSTRACT

Discrete orthogonal functions are used in adaptive system identification algorithms. Adaptive filters are realized by forming linear combinations of discrete Legendre, Laguerre, and Jacobi polynomials, and backward prediction-error polynomials from a lattice structure. The adaptive filter weights are updated using the LMS algorithm. FIR and IIR bandpass filters are modeled using the adaptive filters, and performance comparisons are made.

TABLE OF CONTENTS

LIST OF FIGURES

I. INTRODUCTION

A. ADAPTIVE SYSTEM IDENTIFICATION ALGORITHMS

Adaptive FIR and IIR digital filters are widely used in system identification algorithms. The extensive use of IIR adaptive filters has been hindered by several problems: (1) Instabilities encountered in their design; (2) Local minimums in the mean square error surface; and (3) Slow convergence rates of the adaptive filter weights [Ref. 1]. Although FIR adaptive filters do not suffer from the problems that ail IIR adaptive filters, they typically require much larger orders to effectively model a given system.

Using orthogonal functions in system identification algorithms was first employed by Lee in 1932 [Ref. 2]. Adaptive filters based on orthogonal functions have several advantages, namely, they are always stable and have an infinite impulse response, making their use ideal for modeling systems with long impulse responses. Additionally, the FIR adaptive filter weights converge faster than those associated with an IIR filter structure due to the unimodal mean square error surface. The filter is realized by forming linear combinations of discrete orthogonal functions, which are weighted to minimize the mean square error of the approximation.

B. THESIS OUTLINE

A brief description of contents of the remaining chapters follows. Chapter II introduces the theory of orthogonal functions and their use in modeling linear systems. The definitions of the orthogonal functions included in this research are given. Chapter III develops the actual adaptive filter model that utilizes the orthogonal functions for system identification. The derivation of the filter weights is discussed in detail; significant development of the discrete orthogonal functions is shown. Chapter IV presents the simulation results of the various filters developed in chapter III and comparisons are made between Legendre, Laguerre, Jacobi, and backward prediction-error adaptive digital filters. Chapter V presents conclusions including limitations of the orthogonal polynomial filters and recommendations for further research.

II. ORTHOGONAL FUNCTIONS

A. THEORY OF ORTHOGONAL FUNCTIONS

Let $\{w_1(\tau), w_2(\tau), \ldots\}$ denote a set of real and continuous functions. Then the system of functions is said to be *orthogonal* in the range (a,b) if

$$\int_a^b w_m(\tau)\, w_n(\tau)\; d\tau = \begin{cases} 0 & \text{for } m \neq n \\ k_n^2 & \text{for } m = n\,, \end{cases} \tag{1}$$

where k_n is called the norm of the corresponding function [Ref. 3]. The orthogonal set $\{w_n(\tau)\}$ is considered *complete* if either of the following conditions is true [Ref. 2]:

(1) There exists no function $x(\tau)$ with

$$\int_a^b x^2(\tau)\; d\tau < \infty \tag{2}$$

such that

$$\int_a^b x(\tau)\, w_n(\tau)\; d\tau = 0 \quad , \quad n = 0,\, 1,\, 2,\, \ldots \quad . \tag{3}$$

(2) For any piecewise continuous function $h(\tau)$ with

$$\int_a^b h^2(\tau)\; d\tau < \infty \tag{4}$$

and an $\varepsilon > 0$, there exists an integer N and a polynomial

$$\sum_{n=0}^{N} c_n w_n(\tau) \tag{5}$$

such that

$$\int_a^b \left| h(\tau) - \sum_{n=0}^{N} c_n w_n(\tau) \right|^2 d\tau < \epsilon \quad . \tag{6}$$

Any stable causal system, $h(\tau)$, satisfies (4) in the interval $[0, \infty)$ and can be represented by a complete set of orthogonal functions [Ref. 2]. Letting $\{w_n(\tau)\}$ represent a complete set of orthogonal functions in the interval $[0,\infty)$, then

$$h(\tau) = \begin{cases} \sum_{n=0}^{\infty} c_n w_n(\tau) & for\ 0 \le \tau \le \infty \\ 0 & else \ , \end{cases} \tag{7}$$

where the c_n represent the expansion coefficients.

Albeit impossible to form an infinite sum of orthogonal functions, it is practical to form a finite sum of orthogonal functions with an accompanying error, ϵ, as given by (6). It is therefore possible to form an approximate synthesis of a linear system, $\hat{h}(\tau)$, by forming finite linear combinations of orthogonal functions:

$$\hat{h}(\tau) = \begin{cases} \sum_{n=0}^{N} c_n w_n(\tau) & for\ 0 \le \tau \le \infty \\ 0 & else \ . \end{cases} \tag{8}$$

This, in itself, is not particularly significant, for there are other families of functions that are not orthogonal

that satisfy (8). A Taylor series expansion, for instance, is never orthogonal on any interval, but it is often effective when approximating functions [Ref. 3]. Orthogonal functions, however, have several desirable characteristics that make their use uniquely advantageous when synthesizing linear systems.

Whereas (8) is based on a continuous set of orthogonal functions, it is also possible to form an approximate synthesis of a discrete linear system, $\hat{h}(k)$, such that

$$\hat{h}(k) = \sum_{n=0}^{N} c_n w_n(k) \ , \ k = 0, 1, 2, \ldots , \tag{9}$$

where $\{w_n(k)\}$ represents a complete set of discrete orthogonal functions, and the expansion coefficients, c_n, are chosen to minimize the mean square value of the approximation error.

B. CLASSICAL ORTHOGONAL POLYNOMIALS

The *classical* orthogonal polynomials form a subset of orthogonal functions. The three classical orthogonal polynomial families that are included in this research are the Jacobi, Legendre, and Laguerre. Other families such as the Chebyshev and Hermite were found to be unsuitable for linear system synthesis using the methods described herein.

The Legendre polynomials, denoted by $p_n(\tau)$ [Ref. 4], are orthogonal on the interval [-1,1], and, in the form of (1), their norm is given by

$$\int_{-1}^{1} | p_n(\tau) |^2 \, d\tau = \frac{2}{2n + 1} \quad . \tag{10}$$

The Jacobi polynomials, denoted by $p_n^{(\alpha,\beta)}(\tau)$ [Ref. 4], are also orthogonal on the interval [-1,1] and their norm is given by

$$\int_{-1}^{1} (1 - \tau)^{\alpha} (1 + \tau)^{\beta} \, | p_n^{(\alpha,\beta)}(\tau) |^2 \, d\tau = \frac{2^{\alpha+\beta+1} \, \Gamma(n+\alpha+1) \, \Gamma(n+\beta+1)}{(2n+\alpha+\beta+1) \; n! \; \Gamma(n+\alpha+\beta+1)} \quad . \tag{11}$$

Note that if a substitution is made into (11) with both $\alpha = 0$ and $\beta = 0$, the result is equivalent to (10). Therefore, the Legendre polynomials form a subset of the Jacobi polynomials. Given the orthogonality interval of the Legendre and Jacobi polynomials, [-1,1], it is not possible to make a direct substitution of the polynomials into (7). The desired orthogonality interval for synthesizing causal linear systems is $[0,\infty)$; thus, the orthogonality interval of the Jacobi and Legendre polynomials must be *shifted* by means of a change of variables. Details of this process are discussed in chapter III.

The Laguerre polynomials, $l_n(\tau)$ [Ref. 2], are orthogonal in the interval $[0,\infty)$, making them more readily applied to the synthesis of linear systems than Jacobi and Legendre polynomials. Their norm is given by

$$\int_0^{\infty} | \, l_n(\tau) \, |^2 \, d\tau = 1 \quad . \tag{12}$$

C. ORTHOGONALITY OF THE LATTICE PREDICTOR

Assuming a stationary input of random variables, each stage of the lattice predictor is known to produce a sequence of uncorrelated random variables in the form of the backward prediction errors, $\{b_0(k),\ b_1(k),\ \ldots\}$ [Ref. 3]. These backward errors are orthogonal in the range $[0,\infty)$ and, therefore, well suited for synthesizing discrete linear systems of the form

$$h(k) = \begin{cases} \sum_{n=0}^{\infty} c_n b_n(k) & \text{for } 0 \le k < \infty \\ 0 & \text{else} \end{cases} \quad . \tag{13}$$

The lattice filter structure is a manifestation of the Gram-Schmidt orthogonalization procedure insofar as the generation of the backward errors is concerned. The forward prediction errors associated with prediction-error filters are also produced by each stage of the lattice filter. However, their application to linear system synthesis is not germane, for the forward errors are correlated and, therefore, not orthogonal [Ref. 3].

III. SYSTEM IDENTIFICATION AND MODELING

A. ORTHOGONAL ADAPTIVE FILTER MODEL

Consider a system where x(k) denotes an input to both a causal linear system, h(k), and an adaptive filter model, $\hat{h}(k)$ as shown in Figure 3.1. Let d(k) be the desired output of the system and y(k) be the output of the adaptive filter.

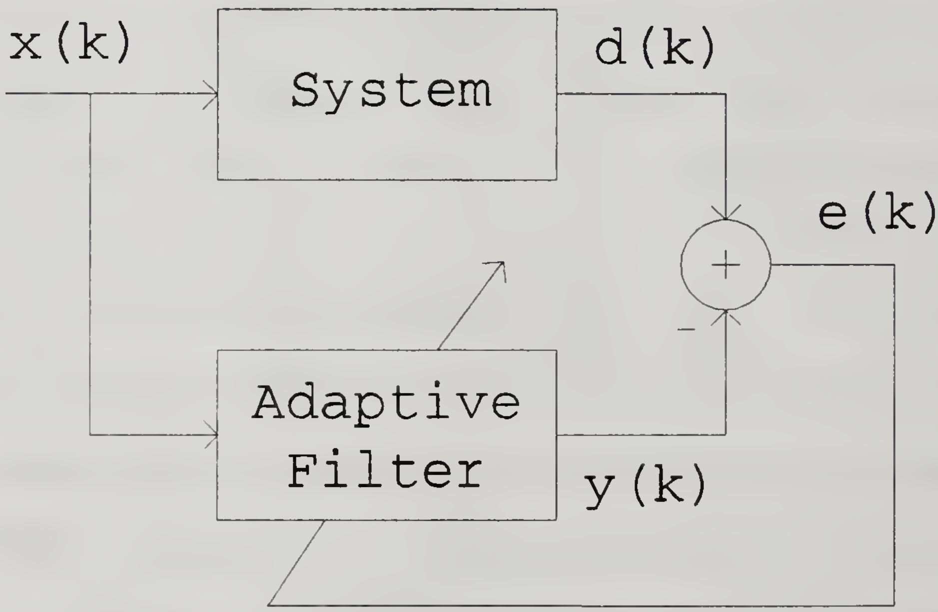

Figure 3.1: System identification configuration

Following the derivation in [Ref. 1], the output error is given by

$$e(k) = d(k) - y(k)$$
$$= d(k) - \sum_{i=-\infty}^{k} x(i)\ \hat{h}(k - i) \quad . \qquad (14)$$

From (9), we write

$$e(k) = d(k) - \sum_{i=-\infty}^{k} x(i) \sum_{n=0}^{N} c_n w_n(k - i) \ , \qquad (15)$$

where $\{w_n(k)\}$ represents a complete set of discrete orthogonal functions, and the c_n are the expansion coefficients. Rearranging the summations in (15) yields

$$e(k) = d(k) - \sum_{n=0}^{N} c_n \sum_{i=-\infty}^{k} x(i) w_n(k - i) \quad . \qquad (16)$$

And therefore,

$$e(k) = d(k) - \sum_{n=0}^{N} c_n u_n(k) = d(k) - \boldsymbol{c}^T \boldsymbol{u}(k) \ , \qquad (17)$$

where

$$\boldsymbol{c} = [c_0,\ c_1,\ \cdots,\ c_N]^T \ , \qquad (18)$$

$$\boldsymbol{u}(k) = [u_0(k),\ u_1(k),\ \cdots,\ u_N(k)]^T \ , \qquad (19)$$

and

$$u_n(k) = \sum_{i=-\infty}^{k} x(i) w_n(k - i) \quad . \qquad (20)$$

Figure 3.2 depicts the generic orthogonal function model. We consider each $Q_n(z)$ to be a black box that has two outputs when excited by an impulse: One is an orthogonal function

$u_n(k)$, and the other is a connection to the next black box. For the purposes of this thesis, we limit our investigation to Legendre, Jacobi, Laguerre, and backward prediction-error functions to model systems.

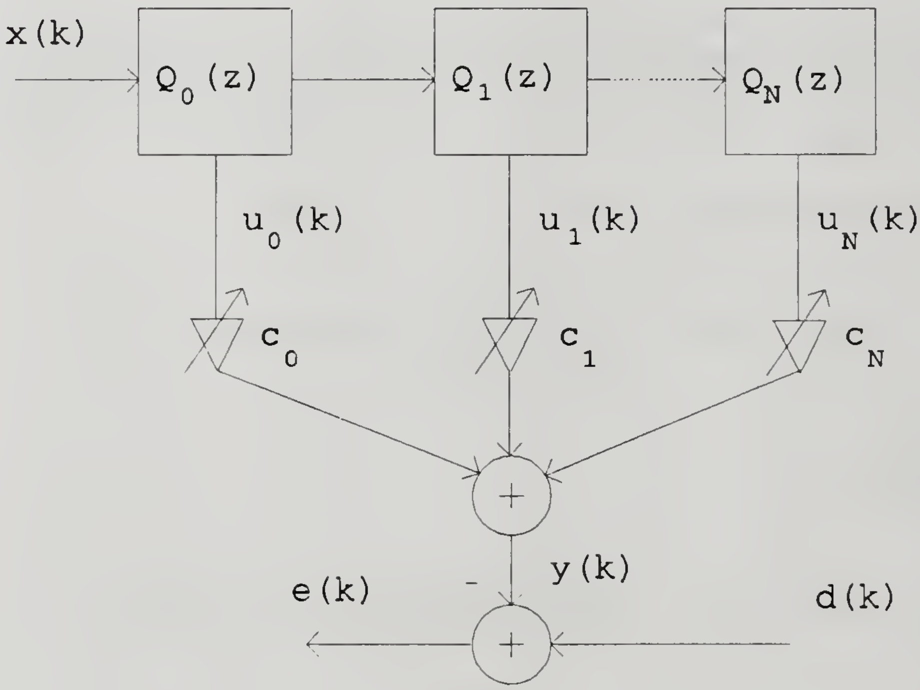

Figure 3.2: Generic adaptive filter configuration for orthogonal functions.

We wish to find a set of expansion coefficients to minimize the mean square error of (17). Using the LMS algorithm, we take $e^2(k)$ to be an estimate of the instantaneous mean square error [Ref. 6]. To obtain the minimum mean square error, we find the corresponding gradient estimate by taking the derivatives of $e^2(k)$ with respect to the expansion coefficients:

$$\nabla(k) = \frac{\partial e^2(k)}{\partial \boldsymbol{c}} = 2e(k)\frac{\partial e(k)}{\partial \boldsymbol{c}} = -2e(k)\boldsymbol{u}(k) \quad . \tag{21}$$

Applying the method of steepest descent, the LMS algorithm updates the expansion coefficients using

$$\boldsymbol{c}(k+1) = \boldsymbol{c}(k) + \mu(-\nabla) \; , \tag{22}$$

where μ is a constant that regulates the convergence rate. Substituting (21) into (22), we obtain the LMS algorithm:

$$\boldsymbol{c}(k+1) = \boldsymbol{c}(k) + 2\mu e(k)\boldsymbol{u}(k) \quad . \tag{23}$$

The expansion coefficients of vector **c** converge in the mean when [Ref. 6]

$$0 < \mu \le \frac{1}{|\boldsymbol{u}(k)|^2} \quad . \tag{24}$$

All simulations in this thesis set μ according to the range specified in (24). The expansion coefficients are updated after each iteration in accordance with (23). Convergence rates for the expansion coefficients vary depending on model type and order.

A. LEGENDRE POLYNOMIALS

Recall from chapter II that continuous Legendre polynomials form a complete set of orthogonal functions in the range [-1,1] and are defined as

$$\int_{-1}^{1} p_m(\tau)p_n(\tau)\ d\tau = \begin{cases} \dfrac{2}{2n+1} & for\ m = n \\ 0 & for\ m \neq n \end{cases} \quad . \tag{25}$$

Following the derivation given by Lee [Ref. 2], a change of variable is made to transform the orthogonality range for the Legendre polynomials from $[-1,1]$ to $[0,\infty)$, which is the desired range to correspond to the causal time axis. Letting the first change of variable be

$$\tau = 2y - 1 \tag{26}$$

causes (25) to become

$$2\int_{0}^{1} p_m(2y-1)p_n(2y-1)\ dy = \begin{cases} \dfrac{2}{2n+1} & for\ m = n \\ 0 & for\ m \neq n \end{cases} \quad . \tag{27}$$

Letting the second change of variable in (27) be

$$y = e^{-Ct}\ , \tag{28}$$

where C is any positive real constant, yields

$$\begin{aligned} &\int_{0}^{\infty} Ce^{-Ct}p_m(2Ce^{-Ct}-1)p_n(2e^{-Ct}-1)\ dt \\ &\qquad = \begin{cases} \dfrac{1}{2n+1} & for\ m = n \\ 0 & for\ m \neq n \end{cases} \quad . \end{aligned} \tag{29}$$

Now, defining

$$v_n(t) = \sqrt{C}\ e^{(-C/2)t}\ p_n(2e^{-Ct}-1) \tag{30}$$

and substituting into (29) gives

$$\int_0^\infty v_m(t)\, v_n(t)\; dt = \begin{cases} \dfrac{1}{2n+1} & \text{for } m = n \\ 0 & \text{for } m \neq n \end{cases} \quad . \tag{31}$$

Since (31) satisfies the definition of orthogonality given by (1), then the set $\{v_n(t)\}$, a shifted version of the Legendre polynomials, is an orthogonal set defined in the desired range $[0,\infty)$. From (7), $h(\tau)$ can now be represented by the series given by

$$h(\tau) = \begin{cases} \sum\limits_{n=0}^{\infty} c_n v_n(\tau) & \text{for } 0 \le \tau \le \infty \\ 0 & \text{else} \,, \end{cases} \tag{32}$$

where, again, c_n are the expansion coefficients chosen to minimize the mean square error, and $\{v_n(t)\}$ is the orthogonal polynomial set based on the shifted Legendre polynomials.

In order to realize a digital network, it is necessary to generate the shifted discrete Legendre polynomials based on the continuous set $\{v_n(t)\}$. It is convenient to note the first few terms of the Legendre polynomials [Ref. 7], $\{p_n(\tau)\}$:

$$\begin{aligned} p_0(\tau) &= 1 \\ p_1(\tau) &= \tau \\ p_2(\tau) &= \frac{3}{2}\tau^2 - \frac{1}{2} \quad . \end{aligned} \tag{33}$$

Substituting (33) into (30), the first few terms of $\{v_n(t)\}$ are

$$\begin{aligned} v_0(t) &= \sqrt{C}e^{-(C/2)t} \\ v_1(t) &= \sqrt{C}(-1 + 2e^{-Ct})\, e^{-(C/2)t} \\ v_2(t) &= \sqrt{C}(1 - 6e^{-Ct} + 6e^{-2Ct})\, e^{-(C/2)t} \quad . \end{aligned} \tag{34}$$

Taking the Laplace transform of (34) generates the first few terms of the shifted Legendre polynomials in the frequency domain, which are given by

$$
\begin{aligned}
V_0(s) &= \sqrt{C}\,\frac{1}{(s + C/2)} \\
V_1(s) &= \sqrt{3C}\,\frac{(s - C/2)}{(s + C/2)(s + 3C/2)} \\
V_2(s) &= \sqrt{5C}\,\frac{(s - C/2)(s - 3C/2)}{(s + C/2)(s + 3C/2)(s + 5C/2)} \quad .
\end{aligned}
\tag{35}
$$

From (35), it is apparent that the general expression for $\{V_n(s)\}$ is

$$V_n(s) = A_n\,\frac{(s - C/2)(s - 3C/2)\cdots(s - (2n - 1)C/2)}{(s + C/2)(s + 3C/2)\cdots(s + (2n + 1)C/2)} \quad , \tag{36}$$

for n = 0, 1, 2, ..., where

$$A_n = \sqrt{(2n + 1)C} \quad . \tag{37}$$

Of the several techniques available for performing an analog to digital filter transformation, the matched Z-transform is best suited for our purposes. This technique preserves the ability to express the digital frequency terms, $\{V_n(z)\}$, in a closed form as in (36) and, more importantly, allows for a filter structure that is easily synthesized. The matched Z-transform maps the poles and zeros of $\{V_n(s)\}$ into those of $\{V_n(z)\}$ through the substitution

$$(s + a) \rightarrow (1 - e^{-aT} z^{-1}) \quad , \tag{38}$$

where T represents the sampling period of the discrete-time filter [Ref. 8].

Accordingly, performing an analog to digital transformation on (36) yields the first three terms of $\{V_n(z)\}$:

$$\begin{aligned} V_0(z) &= \frac{\sqrt{C}}{1 - e^{-C/2}z^{-1}} \\ V_1(Z) &= \frac{\sqrt{3C}\,(1 - e^{C/2}z^{-1})}{(1 - e^{-C/2}z^{-1})(1 - e^{-3C/2}z^{-1})} \\ V_2(z) &= \frac{\sqrt{5C}\,(1 - e^{C/2}z^{-1})(1 - e^{3C/2}z^{-1})}{(1-e^{-C/2}z^{-1})(1-e^{-3C/2}z^{-1})(1-e^{-5C/2}z^{-1})} \quad . \end{aligned} \tag{39}$$

As noted, the matched Z-transform preserves the ability to express the terms of $\{V_n(z)\}$ in the closed form

$$V_n(z) = \frac{A_n\,(1 - e^{C/2}z^{-1})(1 - e^{3C/2}z^{-1})\cdots(1 - e^{(2n-1)C/2}z^{-1})}{(1 - e^{-C/2}z^{-1})(1 - e^{-3C/2}z^{-1})\cdots(1 - e^{-(2n+1)C/2}z^{-1})} \ , \tag{40}$$

for n = 0, 1, 2, ..., where

$$A_n = \sqrt{(2n + 1)C} \ . \tag{41}$$

The synthesis of a discrete system using shifted Legendre polynomials is accomplished by generating each Legendre polynomial using its corresponding transfer function in (40). Accordingly, substituting each $Q_n(z)$ shown in Figure 3.3 into the filter structure shown in Figure 3.2 provides the necessary filter structure to generate each polynomial.

C. JACOBI POLYNOMIALS

The Legendre polynomials form a subset of the much larger class of Jacobi polynomials. Like the Legendre polynomials, the Jacobi polynomials form a complete set of orthogonal

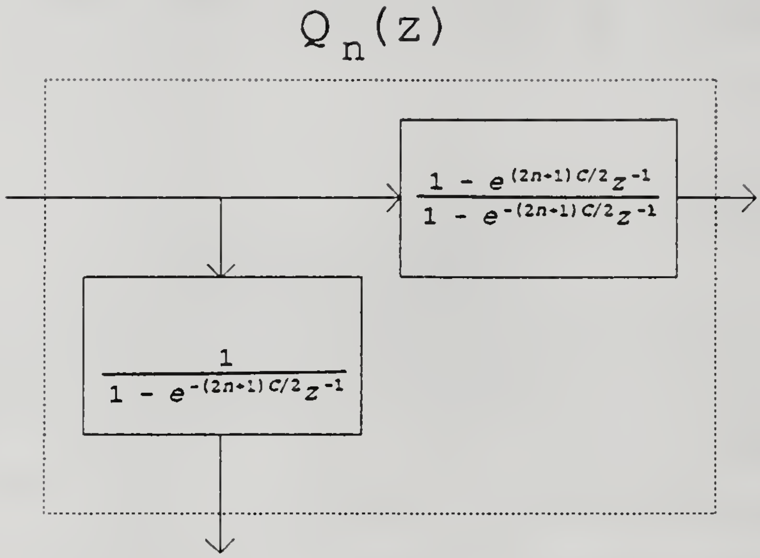

Figure 3.3: Module of Legendre polynomial adaptive filter.

functions in the range [-1,1] and, therefore, require a transformation to shift the orthogonality range to $[0,\infty)$.

The Jacobi polynomials, $\{p_n^{(\alpha,\beta)}(\tau)\}$, are defined as

$$\int_{-1}^{1} (1-\tau)^{\alpha}(1+\tau)^{\beta}\, p_m^{(\alpha,\beta)}(\tau)\, p_n^{(\alpha,\beta)}(\tau)\; d\tau = \begin{cases} k_n^2 & for\ m = n \\ 0 & for\ m \neq n\ , \end{cases} \tag{42}$$

where

$$k_n^2 = \frac{2^{\alpha+\beta+1}\,\Gamma(n+\alpha+\beta+1)\;\Gamma(n+\beta+1)}{(2n+\alpha+\beta+1)\; n!\;\Gamma(n+\alpha+\beta+1)} \quad . \tag{43}$$

Letting the first change of variable be

$$\tau = 2y - 1 \tag{44}$$

causes (42) to become

$$2 \int_0^1 (2 - 2y)^{\alpha} (2y)^{\beta}\, p_m^{(\alpha,\beta)}(2y - 1)\, p_n^{(\alpha,\beta)}(2y - 1)\, dy = \begin{cases} k_n^2 & if\ m = n \\ 0 & if\ m \neq n \end{cases} . \tag{45}$$

Letting the second change of variable in (45) be

$$y = e^{-Ct} , \tag{46}$$

where C is any positive real constant, yields

$$2 \int_0^{\infty} Ce^{-Ct}(2 - 2e^{-Ct})^{\alpha} (2e^{-Ct})^{\beta}\, p_m^{(\alpha,\beta)}(2e^{-Ct} - 1)\; X\, p_n^{(\alpha,\beta)}(2e^{-Ct} - 1)\, dt = \begin{cases} k_n^2 & for\ m = n \\ 0 & for\ m \neq n . \end{cases} \tag{47}$$

Now, defining

$$v_n^{(\alpha,\beta)}(t) = \sqrt{2C}\, 2^{\beta/2} (2 - 2^{-Ct})^{\alpha/2}\, e^{-C(\beta+1)t/2}\, p_n^{(\alpha,\beta)}(2e^{-Ct} - 1) , \tag{48}$$

and substituting into (47) gives

$$\int_0^{\infty} v_m^{(\alpha,\beta)}(t)\, v_n^{(\alpha,\beta)}(t)\, dt = \begin{cases} k_n^2 & for\ m = n \\ 0 & for\ m \neq n , \end{cases} \tag{49}$$

where k_n^2 is defined as in (43). Since (49) satisfies the definition of orthogonality given in (1), the shifted Jacobi polynomials, $\{v_n^{(\alpha,\beta)}(t)\}$, form an orthogonal set of polynomials in the range $[0,\infty)$. From (7) , any causal system, h(t), may be represented by

$$h(t) = \begin{cases} \sum_{n=0}^{\infty} c_n v_n^{(\alpha,\beta)}(t) & \textit{for } 0 \le t \le \infty \\ 0 & \textit{else} \end{cases} \quad . \tag{50}$$

In order to synthesize discrete linear systems, it is necessary to generate the shifted discrete Jacobi polynomials. Let the derivation of the shifted discrete Jacobi polynomials begin by listing the first few terms of the continuous Jacobi polynomials:

$$\begin{aligned} p_0^{(\alpha,\beta)}(t) &= 1 \\ p_1^{(\alpha,\beta)}(t) &= (\alpha+1) + \frac{1}{2}(\alpha+\beta+2)(t-1) \\ p_2^{(\alpha,\beta)}(t) &= \frac{1}{2}(\alpha+1)(\alpha+2) + \frac{1}{2}(\alpha+\beta+3)(\alpha+2)(t-1) \\ &\quad + \frac{1}{8}(\alpha+\beta+3)(\alpha+\beta+4)(t-1)^2) \quad . \end{aligned} \tag{51}$$

Substituting (51) into (48) yields the first few terms of $\{v_n^{(\alpha,\beta)}(t)\}$:

$$\begin{aligned} v_0^{(\alpha,\beta)}(t) &= \sqrt{2C}\, 2^{\beta/2}\, e^{-C(\beta+1)t/2}\, (2 - 2e^{-Ct})^{\alpha/2} \\ v_1^{(\alpha,\beta)}(t) &= \sqrt{2C}\, 2^{\beta/2}\, e^{-C(\beta+1)t/2}\, (2 - 2e^{-Ct})^{\alpha/2} \\ &\quad x\, [(\alpha+1) + \frac{1}{2}(\alpha+\beta+2)(2^{-Ct} - 2)] \\ v_2^{(\alpha,\beta)}(t) &= \sqrt{2C}\, 2^{\beta/2}\, e^{-C(\beta+1)t/2}\, (2 - 2e^{-Ct})^{\alpha/2} \\ &\quad x\, [(\frac{1}{2}(\alpha+1)(\alpha+2) \\ &\quad + \frac{1}{2}(\alpha+\beta+3)((\alpha+2)(2e^{-Ct} - 2) \\ &\quad + \frac{1}{8}(\alpha+\beta+3)(\alpha+\beta+4)(2e^{-Ct} - 2)^2] \quad . \end{aligned} \tag{52}$$

We desire to obtain $\{V_n^{(\alpha,\beta)}(s)\}$, the frequency domain terms of the shifted Jacobi polynomials; however, it is not possible to form a general expression for each term of $\{V_n^{(\alpha,\beta)}(s)\}$ because the Laplace transform of (52) yields entirely

different results depending on the value of α. Therefore, without losing the generality, we make the derivation specific to $\alpha = 2.0$ for the remainder of this development.

Substituting $\alpha = 2.0$ into (52) and taking its Laplace transform, we obtain the first few terms of $\{V_n^{(2,\beta)}(s)\}$:

$$
\begin{aligned}
V_0^{(2,\beta)}(s) &= \frac{K_0}{(s + \frac{C(\beta+1)}{2})\,(s + \frac{C(\beta+3)}{2})} \\
V_1^{(2,\beta)}(s) &= \frac{K_1\,(s - (C+1)/6)}{(s + \frac{C(\beta+1)}{2})\,(s + \frac{C(\beta+3)}{2})\,(s + \frac{C(\beta+5)}{2})} \\
V_2^{(2,\beta)}(s) &= \frac{K_2\,[12s^2 - (4\beta-8)Cs - (\beta^2+6\beta+5)C^2]}{(s+\frac{C(\beta+1)}{2})(s+\frac{C(\beta+3)}{2})(s+\frac{C(\beta+5)}{2})(s+\frac{C(\beta+7)}{2})}\,,
\end{aligned}
\tag{53}
$$

where the K_n are constants.

Unlike the Legendre polynomials, there is no apparent closed form expression to represent the terms of $\{V_n^{(2,\beta)}(s)\}$. The denominator terms can be put into a closed form expression, but not the numerator terms. This is an important concern, for it is carried over to the frequency terms of the discrete shifted Jacobi polynomials, $\{V_n^{(2,\beta)}(z)\}$.

Using the matched Z-transform method to perform an analog to digital transformation on the first two terms in (53) produces the first two terms of $\{V_n^{(2,\beta)}(z)\}$:

$$
\begin{aligned}
V_0^{(2,\beta)}(z) &= \frac{K_{00}}{(1 - a_0 z^{-1})(1 - a_1 z^{-1})} \\
V_1^{(2,\beta)}(z) &= \frac{K_{11}\,(1 - e^{(C+1)/6} z^{-1})}{(1 - a_0 z^{-1})(1 - a_1 z^{-1})(1 - a_2 z^{-1})}\,,
\end{aligned}
\tag{54}
$$

where

$$a_n = e^{\frac{-C(\beta+2n+1)}{2}} , \tag{55}$$

and the K_{nn} are constants.

Without a closed form representation of the terms in (54), the problem of generating the discrete Jacobi polynomials is considerably more demanding than that of the Legendre polynomials. Whereas the transfer function needed to generate each of the shifted Legendre polynomials is known (see (40)), the transfer function needed to generate each of the shifted Jacobi polynomials must be explicitly derived. Due to the complexity of the expressions involved , it was necessary to use a symbolic software program called MACSYMA to generate both the terms of $\{v_n^{(2,\beta)}(t)\}$ and $\{V_n^{(2,\beta)}(s)\}$. The terms of $\{V_n^{(2,\beta)}(z)\}$ could then be found by transformation. Unfortunately, only the first 15 terms of $\{V_n^{(\alpha,\beta)}(z)\}$ were derived with $\alpha = 2.0$ and $\alpha = 4.0$ due to the large size of the expressions involved.

D. LAGUERRE POLYNOMIALS

The Laguerre polynomials, $\{l_n(t)\}$, form a complete set of orthogonal functions in the range $[0,\infty)$ and are defined by

$$\int_0^\infty l_m(t)\, l_n(t)\; dt = \begin{cases} 1 & for\ m = n \\ 0 & for\ m \neq n \end{cases} . \tag{56}$$

Laguerre polynomials are defined over the desired orthogonality range, $[0,\infty)$; thus, we can immediately proceed

with the derivation of the discrete Laguerre polynomials that are needed to synthesize discrete linear systems. The first few terms of $\{l_n(t)\}$ are:

$$\begin{aligned} l_0(t) &= \sqrt{2C}\, e^{-Ct} \\ l_1(t) &= \sqrt{2C}\,(2Ct - 1)\, e^{-Ct} \\ l_2(t) &= \sqrt{2C}(2C^2t^2 - 4Ct + 1)\, e^{-Ct}\,, \end{aligned} \tag{57}$$

where C is any positive real constant. Taking the Laplace transform of (57), we have

$$\begin{aligned} L_0(s) &= \sqrt{2C}\,\frac{1}{(s + C)} \\ L_1(s) &= \sqrt{2C}\,\frac{(s - C)}{(s + C)^2} \\ L_2(s) &= \sqrt{2C}\,\frac{(s - C)^2}{(s + C)^3} \quad . \end{aligned} \tag{58}$$

Using the matched Z-transform technique to perform an analog to digital transformation on (58) yields

$$\begin{aligned} L_0(z) &= \sqrt{2C}\,\frac{1}{(1 - e^{-C}z^{-1})} \\ L_1(z) &= \sqrt{2C}\,\frac{(1 - e^{C}z^{-1})}{(1 - e^{-C}z^{-1})^2} \\ L_2(z) &= \sqrt{2C}\,\frac{(1 - e^{C}z^{-1})^2}{(1 - e^{-C}z^{-1})^3} \quad . \end{aligned} \tag{59}$$

The closed form general expression for the terms in (59) is

$$L_n(z) = \sqrt{2C}\,\frac{(1 - e^{C}z^{-1})^n}{(1 - e^{-C}z^{-1})^{n+1}} \quad . \tag{60}$$

Notice that all zeros of (59) are located at $z=e^{C}$, and all poles are located at $z=e^{-C}$. The synthesis of any linear system using Laguerre polynomials is accomplished by generating each Laguerre polynomial using its corresponding transfer function

given in (60). Accordingly, substituting each $Q_n(z)$ in Figure 3.4 into the structure shown in Figure 3.2 provides the necessary filter structure to generate the polynomials.

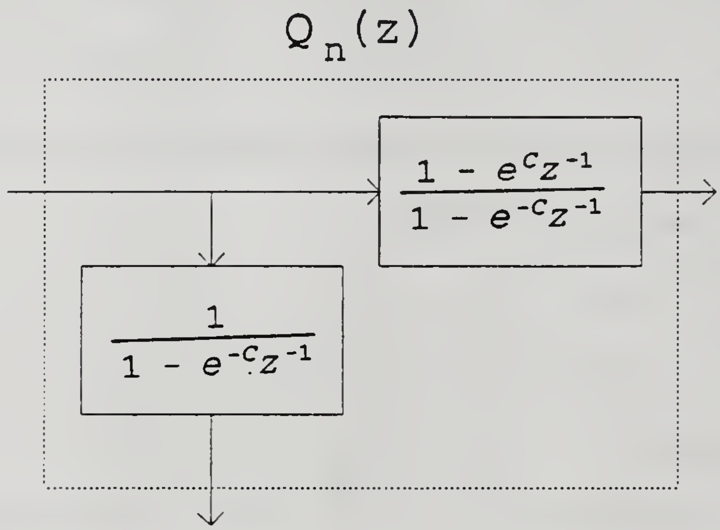

Figure 3.4: Module of Laguerre polynomial adaptive filter.

E. LATTICE PREDICTOR

Let $\{b_0(k),\ b_1(k),\ \ldots,\ b_N(k)\}$ denote the first N+1 backward prediction errors associated with a backward prediction error-filter. If x(k) is a stationary input of random variables to a backward prediction-error filter, it can be shown that the backward errors are orthogonal, that is,

$$E[b_m(k)\ b_n(k)] = \begin{cases} k_n^2 & for\ m = n \\ 0 & for\ m \neq n\ , \end{cases} \qquad \textbf{(61)}$$

where E[] is the expectation operator. Following the proof given by Haykin [Ref. 5], we write

$$b_m(k) = \sum_{l=0}^{m} a_m(m - l)\ x(k - l)\ , \tag{62}$$

where $a_m(l)$, $l = 0, 1, \ldots, m$, are the coefficients of a prediction-error filter of order m. Substituting (62) into (61) yields

$$\begin{aligned} E[b_m(k)\ b_n(k)] &= E\left[\sum_{l=0}^{m} \sum_{p=0}^{n} a_m(m - l)\ a_n(n - p)\ x(k - l)\ x(k - p)\right] \\ &= \sum_{l=0}^{m} \sum_{p=0}^{n} a_m(m - l)\, a_n(n - p)\ E[x(k - l)\ x(k - p)] \\ &= \sum_{l=0}^{m} \sum_{p=0}^{n} a_m(m - l)\ a_n(n - p)\, r_x(p - l)\ , \end{aligned} \tag{63}$$

where $r_x(p-l)$ denotes the correlation function. The normal equations for a backward prediction-error filter are given by [Ref. 5]

$$\sum_{p=0}^{n} a_n(n - p)\ r_x(p - l) = \begin{cases} k_n^2 & \text{for } l = n \\ 0 & \text{for } l \neq n \end{cases} . \tag{64}$$

Substituting (64) into (65), we find

$$E[b_m(k)\ b_n(k)] = \begin{cases} k_n^2 & \text{for } m = n \\ 0 & \text{for } m \neq n \end{cases} . \tag{65}$$

A lattice predictor structure is commonly used to generate the backward errors. The lattice structure shown in Figure 3.5 utilizes the backward errors in an adaptive filter. Let the transfer function needed to generate each $\{b_n(k)\}$ be denoted by

$$A_n(z) = \frac{B_n(z)}{X(z)} \quad . \tag{66}$$

By comparing Figure 3.5 to Figure 3.2 we note that $A_n(z)$ has a definite relationship to $Q_n(z)$. The exact relationship could be found by evaluating each $A_n(z)$ using the signal flow graph in Figure 3.5. Furthermore, it is apparent that $u_n(k) = b_n(k)$, and, thus, the backward prediction-errors could be said to form a family of orthogonal polynomials. The synthesis of a linear system is accomplished by forming linear combinations of the backward prediction-error polynomials.

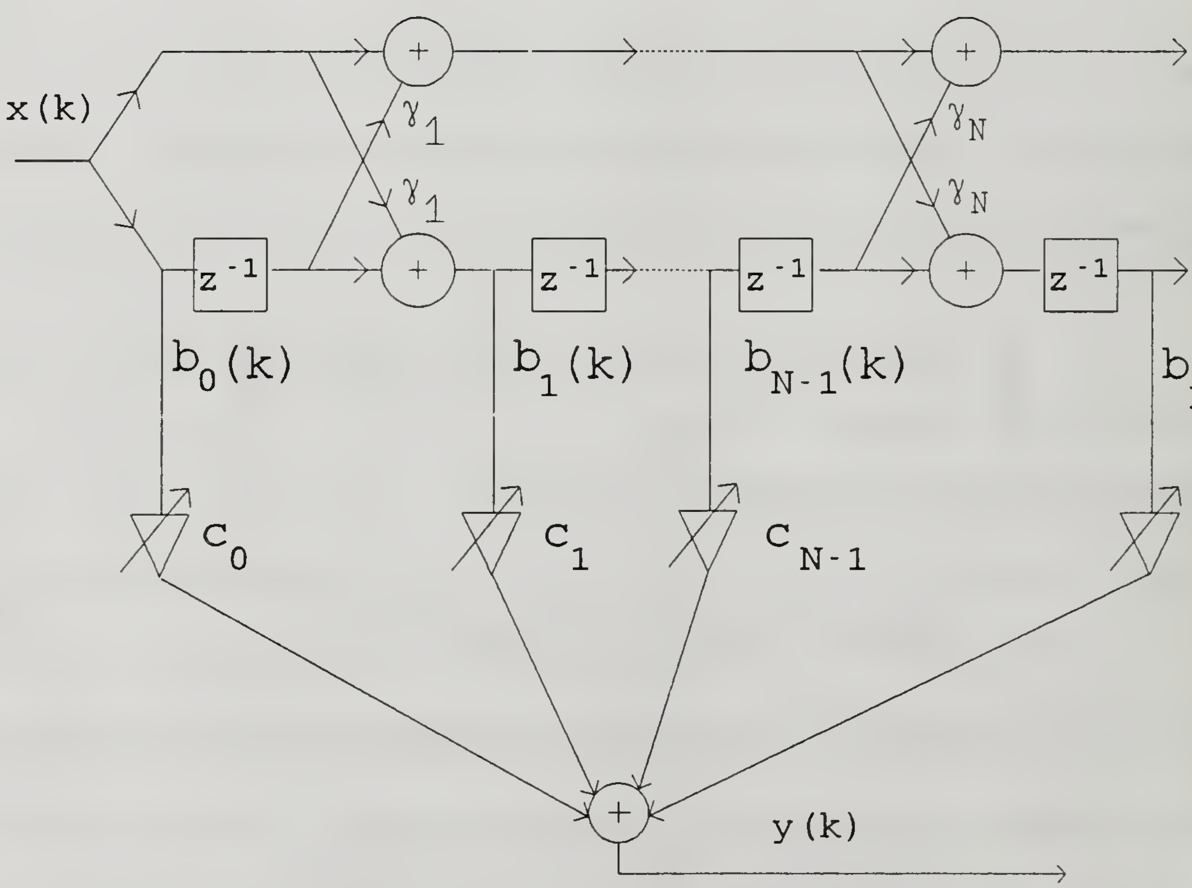

Figure 3.5: Lattice predictor adaptive filter.

IV. SIMULATION RESULTS

A. FIR BANDPASS FILTER SIMULATIONS

For the purpose of evaluating the FIR system modeling performance of the orthogonal polynomials, two bandpass FIR filters were chosen as the systems to be identified. The FIR filters were designed for filter orders of 22 and 71, with cutoff frequencies of 0.1 and 0.2 (fraction of sampling frequency). Each filter was excited by a zero-mean Gaussian white noise sequence with unit variance. The desired output sequence was compared against the output of the orthogonal network, and the error was used to update the expansion coefficients.

Figure 4.1 shows the impulse and frequency response plots of an 18^{th} order Legendre adaptive digital filter (ADF) with $C = 0.11$ when used to model a 22^{nd} order FIR system. Notice that the Legendre filter has an infinite impulse response that allows smaller order Legendre filters to model larger order FIR systems. In this case an 18^{th} order model is used for a 22^{nd} system, which is roughly an 18% savings in terms of the filter order. Figure 4.2 shows a 52^{nd} order Legendre ADF model with $C = 0.075$ used to model a 71^{st} order FIR system, which is approximately a 27% savings in terms of the filter order.

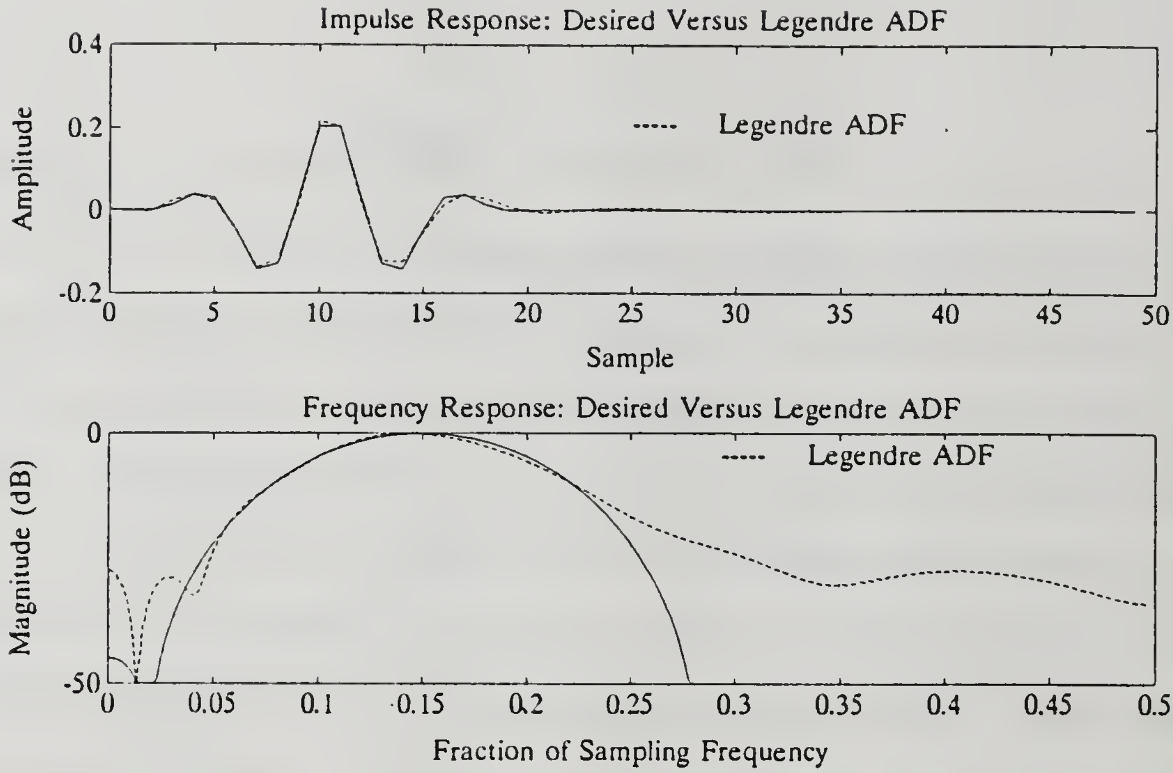

Figure 4.1: Impulse and frequency response of an 18^{th} order Legendre ADF used to model a 22^{nd} order FIR filter.

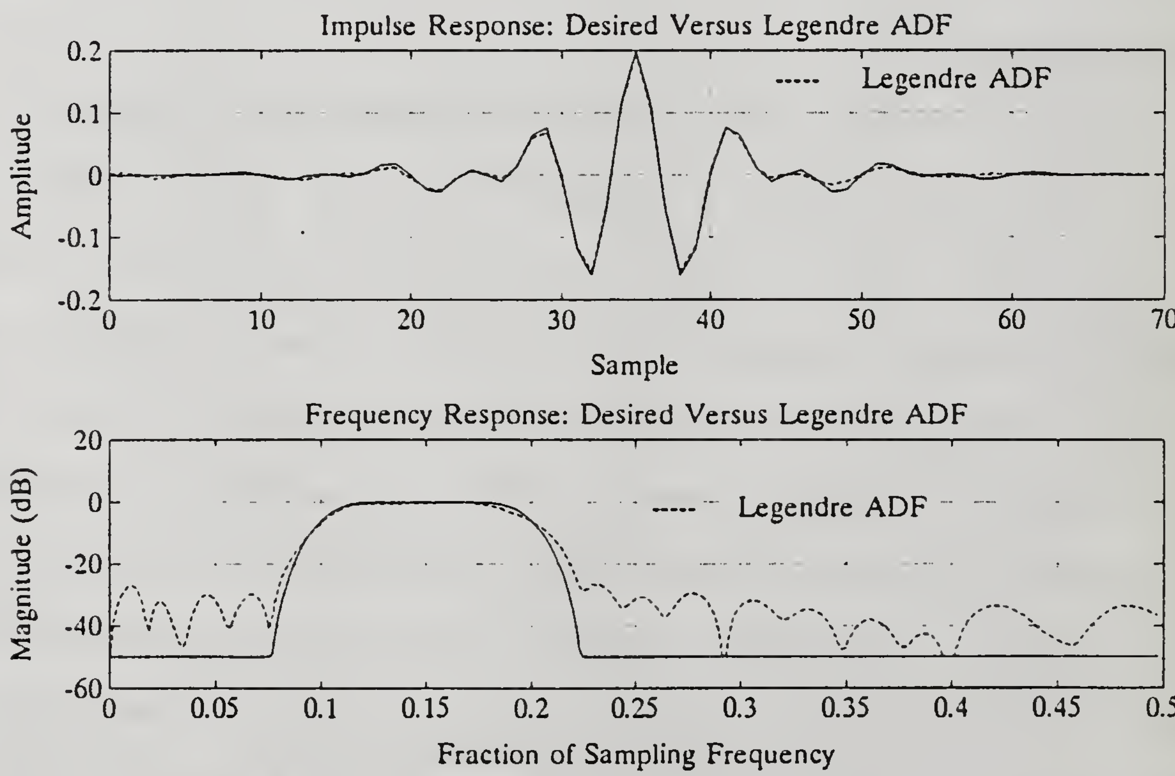

Figure 4.2: Impulse and frequency response of a 52^{nd} order Legendre ADF used to model a 71^{st} order FIR filter.

The convergence of the first six expansion coefficients for the 18^{th} order Legendre ADF and the resulting filter output error are shown in Figure 4.3. Notice that all coefficients converge by the 200^{th} sample, which implies we need an input sequence that has approximately 10 times more samples than the order of the system to be modeled.

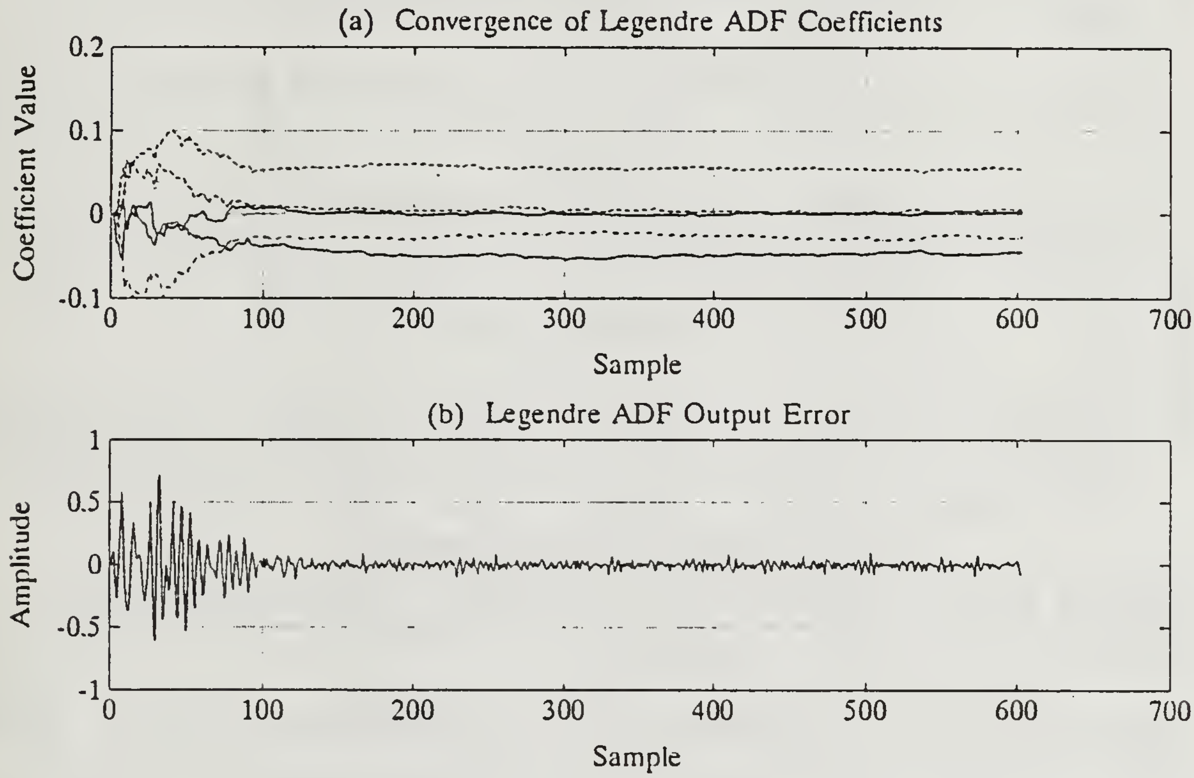

Figure 4.3: (a) Convergence of first six expansion coefficients and (b) Output error for an 18^{th} order Legendre ADF used to model the 22^{nd} order FIR system.

The performance results when an 18^{th} order Laguerre ADF with C = 0.98 is used to model the 22^{nd} order FIR system is shown in Figure 4.4. The fact that the results are comparable to the Legendre ADF is not surprising, for the transfer functions needed to generate both sets of polynomials have the same number of poles and zeros (see (40) and (60)). Likewise,

Figure 4.5 shows the result of a 52^{nd} order model with C = 0.98 used to model the 71^{st} order FIR system.

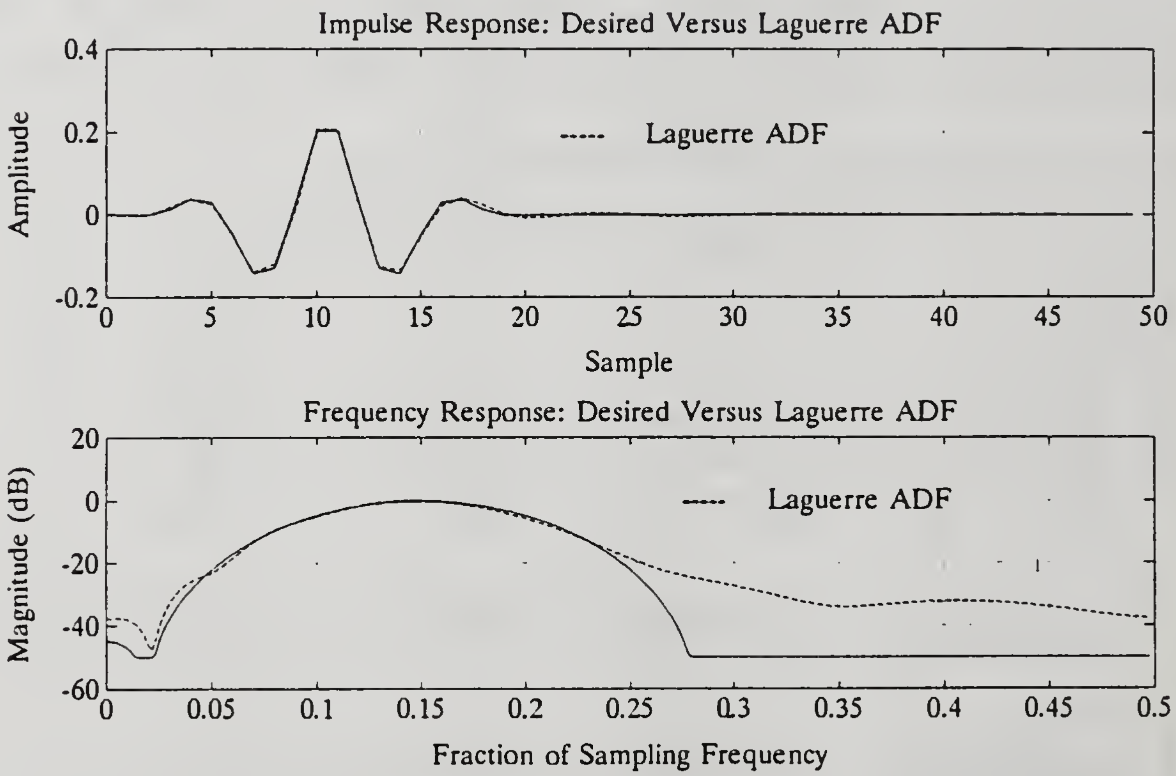

Figure 4.4: Impulse and frequency response of an 18^{tn} order Laguerre ADF used to model a 22^{nd} order FIR filter.

We expect an improvement in performance when using the Jacobi ADF, for the transfer function needed to generate each Jacobi polynomial has more poles than that needed to generate the Legendre and Laguerre polynomials. Exactly how many more poles are realized depends on the value of α. In the case where α = 4.0, two more poles are realized in each transfer function that generates the corresponding Jacobi polynomial. Figure 4.6 shows the results of a 15^{th} order Jacobi simulation with {C, α, β} = {0.002, 4, 1000} when used to model the 22^{nd} order FIR system.

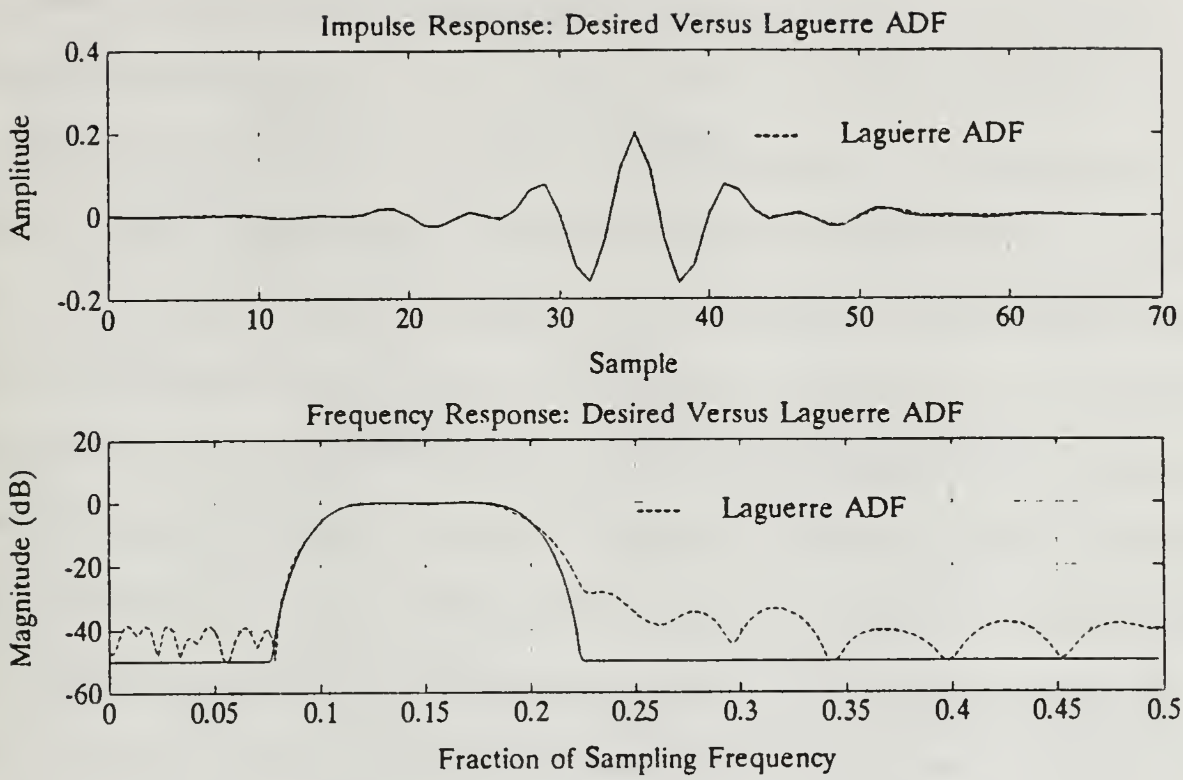

Figure 4.5: Impulse and frequency response of an 52nd order Laguerre ADF used to model a 71st order FIR filter.

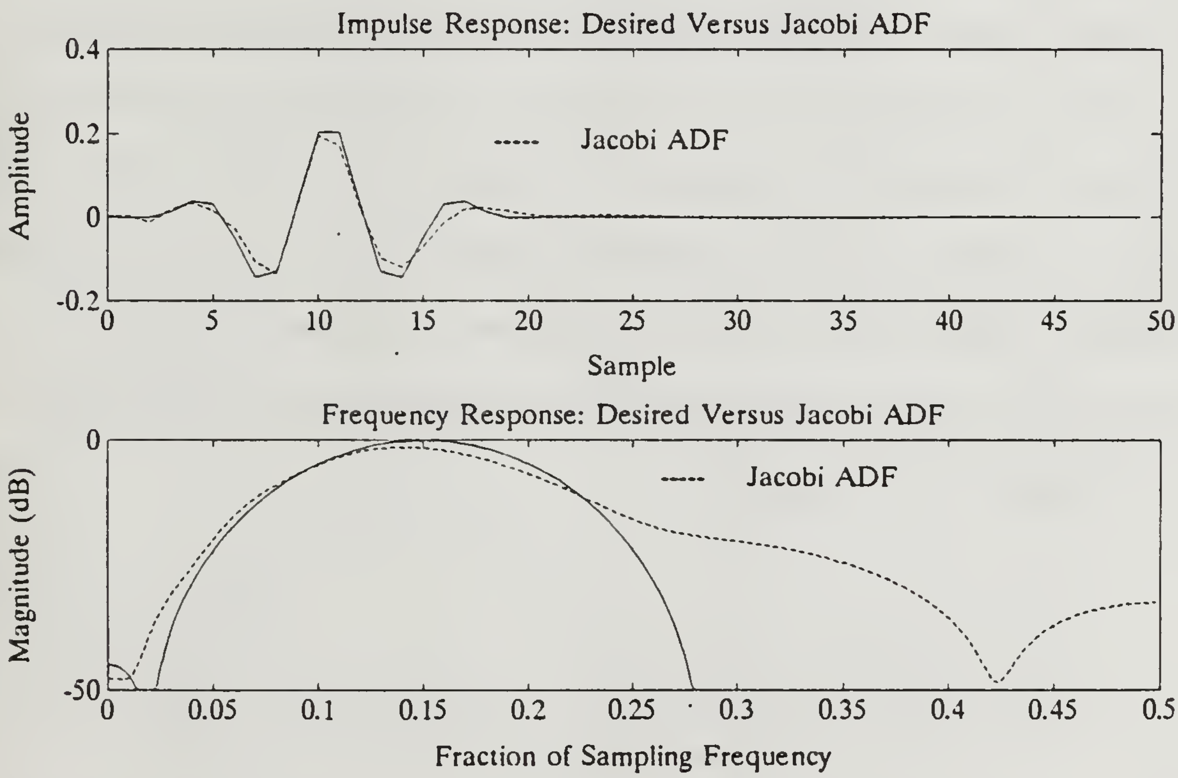

Figure 4.6: Impulse and frequency response of a 15th order Jacobi ADF used to model a 22nd order FIR filter.

Note that the frequency response of the Jacobi model is significantly better in the lower end of the spectrum than in the higher end. This can be attributed to the fact that the frequency response of each Jacobi polynomial has lowpass characteristics; therefore, it is not surprising that a linear combination of Jacobi polynomials would perform better when modeling frequencies in the lower end of the frequency spectrum. The same is true for both the Legendre and Laguerre polynomials. The 71^{st} order FIR system was not modeled using the Jacobi polynomials due to the difficulties encountered in determining the transfer functions of the higher order Jacobi polynomials as discussed in part C of chapter III.

Unlike the Legendre, Laguerre, and Jacobi filters, the lattice filter structure has a finite impulse response. The 15^{th} order Lattice model produces comparable results to that of the classical orthogonal polynomial models when used to model the 22^{nd} order FIR system (see Figure 4.7). Although the lattice filter has a finite impulse response, Figure 4.8 shows that a 52^{nd} order lattice filter is able to effectively model a 71^{st} order FIR system.

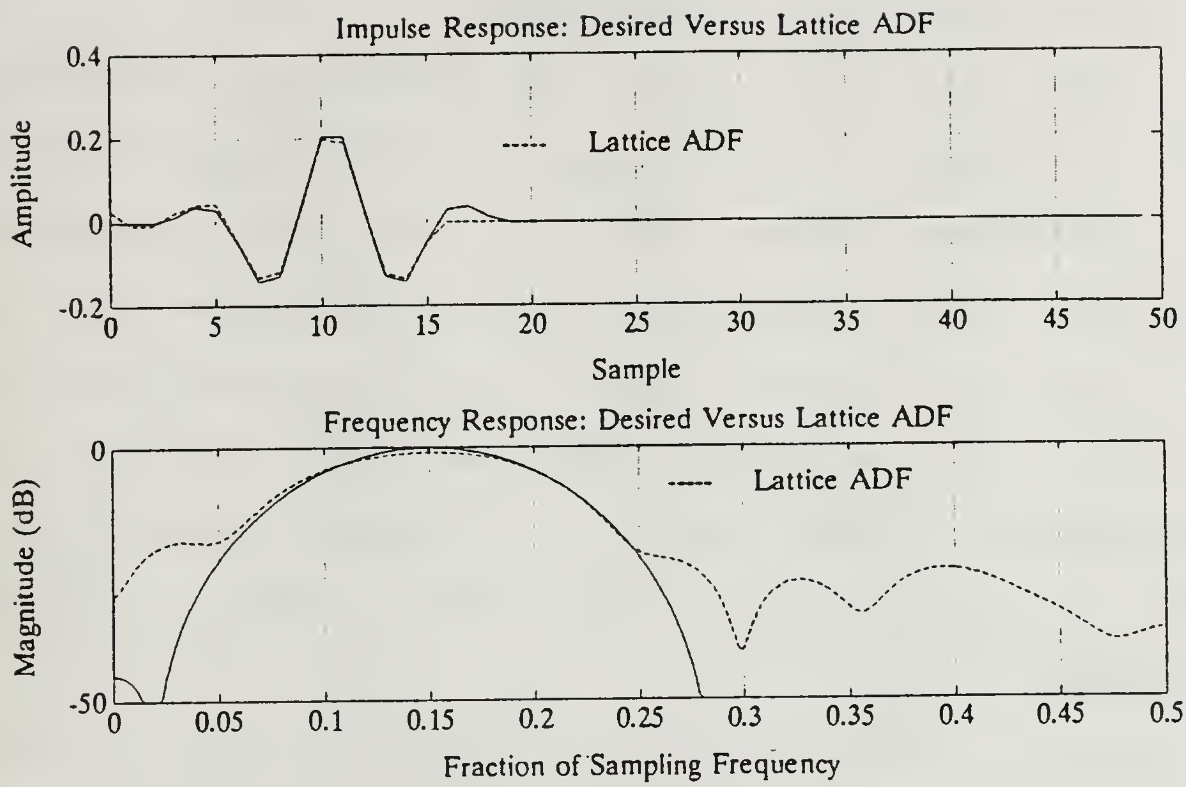

Figure 4.7: Impulse and Frequency response of a 15^{th} order lattice ADF used to model a 22^{nd} order FIR filter.

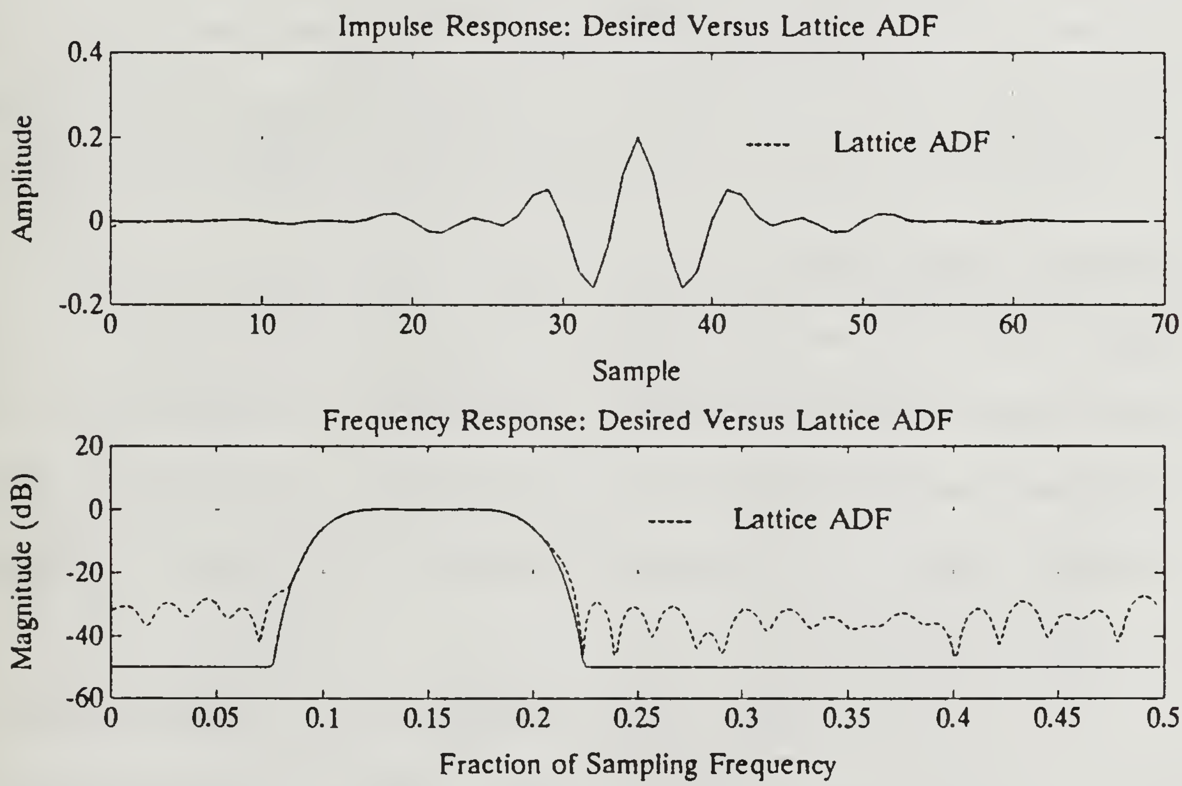

Figure 4.8: Impulse and frequency response of a 52^{nd} order lattice ADF used to model a 71^{st} order FIR filter.

B. IIR BANDPASS FILTER SIMULATIONS

As with the FIR filters, two Butterworth IIR bandpass filters were designed to evaluate the IIR modeling performance of the orthogonal ADFs. The IIR filters were designed for filter orders of 7 and 31, with cutoff frequencies of 0.1 and 0.2. Since an N^{th} order IIR filter has 2N coefficients, our aim is to use orthogonal function adaptive models with approximately 2N coefficients. The results herein show the lowest order ADF models that demonstrated satisfactory performance.

Figure 4.9 shows a 15^{th} order Legendre ADF with C = 0.075 used to model the 7^{th} order Butterworth IIR filter. Notice that the Legendre ADF does not model the low order IIR system as well as the low order FIR system as shown in Figure 4.1. The IIR filters are difficult to model because the structure of the orthogonal adaptive filters more closely resembles a FIR filter (see Figure 3.2).

Figure 4.10 shows a 65^{th} order Legendre ADF with C = 0.055 used to model the 31^{st} order IIR filter. Notice that the impulse response of the IIR filter is still significant beyond 140 samples, but the impulse response of the ADF does a poor job of duplicating it.

As in the FIR case, the Laguerre ADF performance is similar to that of the Legendre ADF. Figure 4.11 shows a 15^{th} order Laguerre ADF with C = 0.98 modeling a 7^{th} order IIR filter. There is marked increase in performance when the

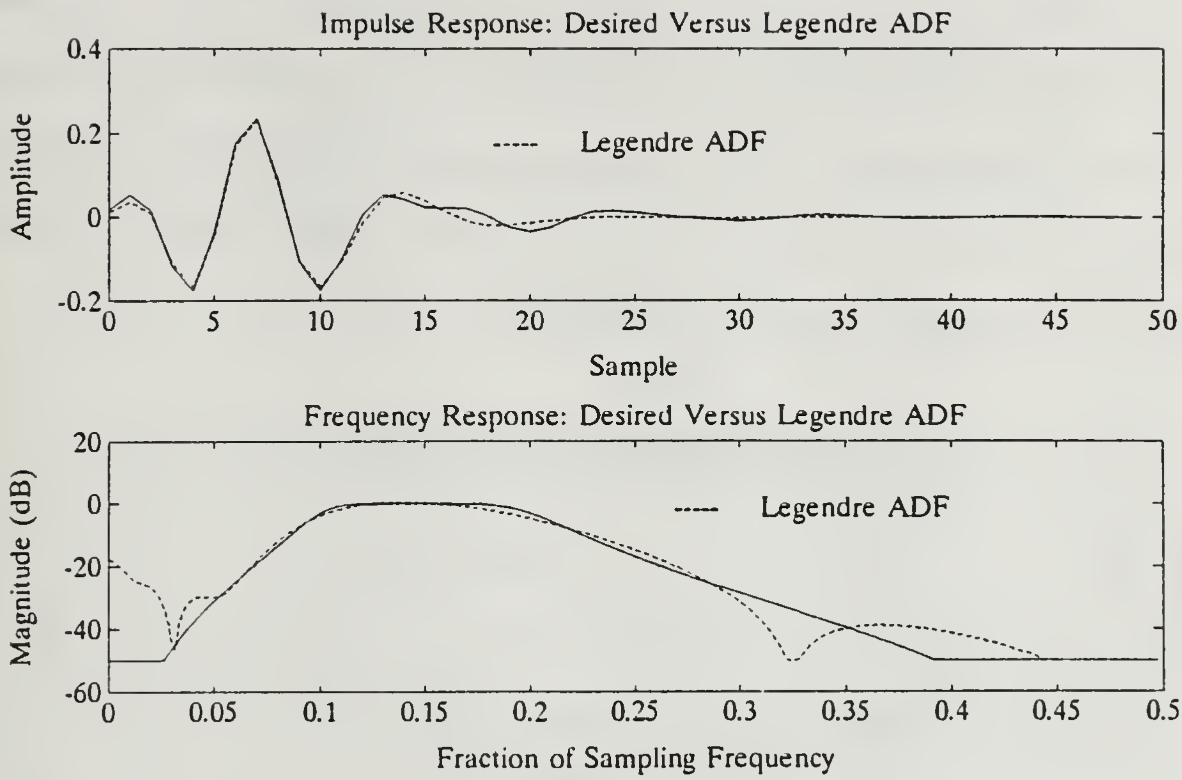

Figure 4.9: Impulse and frequency response of a 15^{th} order Legendre ADF used to model a 7^{th} order IIR filter.

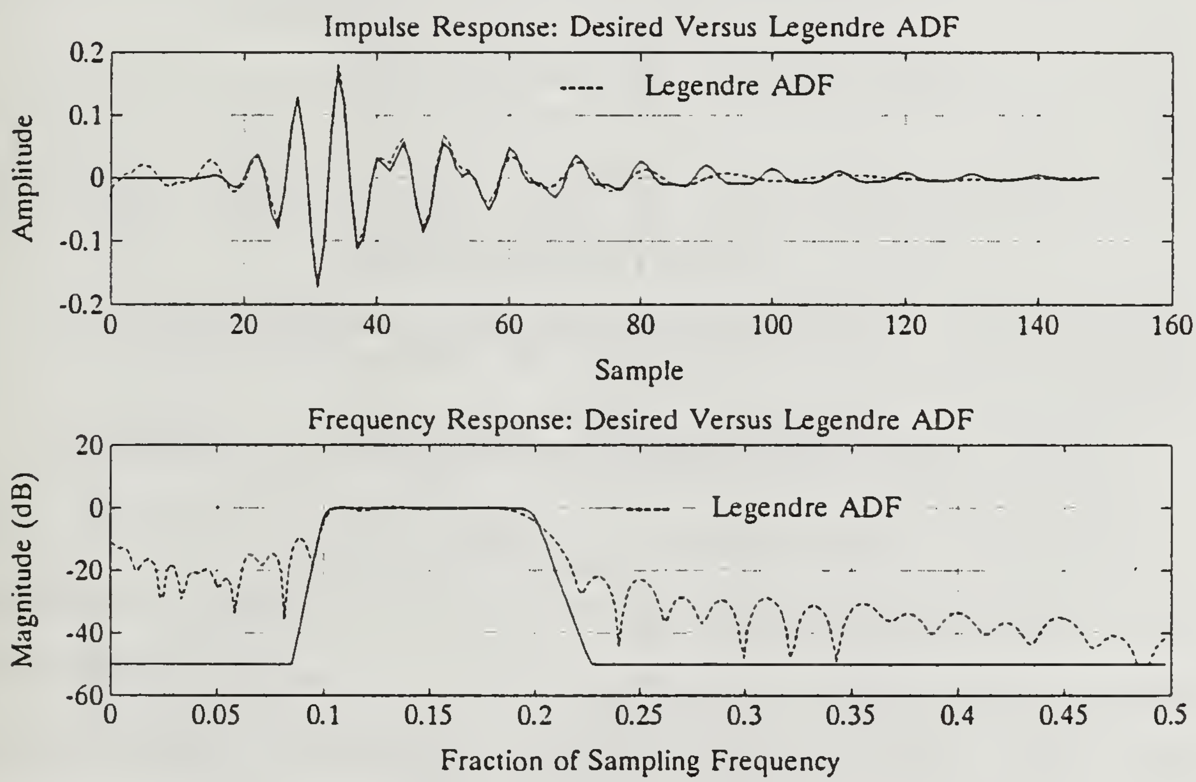

Figure 4.10: Impulse and frequency response of a 65^{th} order Legendre ADF used to model a 31^{st} order IIR filter.

order of the system being identified is increased. Figure 4.12 shows that the 31^{st} order IIR system can be modeled with a 65^{th} order ADF. The Laguerre ADF has better performance characteristics than the Legendre ADF (see Figure 4.10) in both the passband and the stopband.

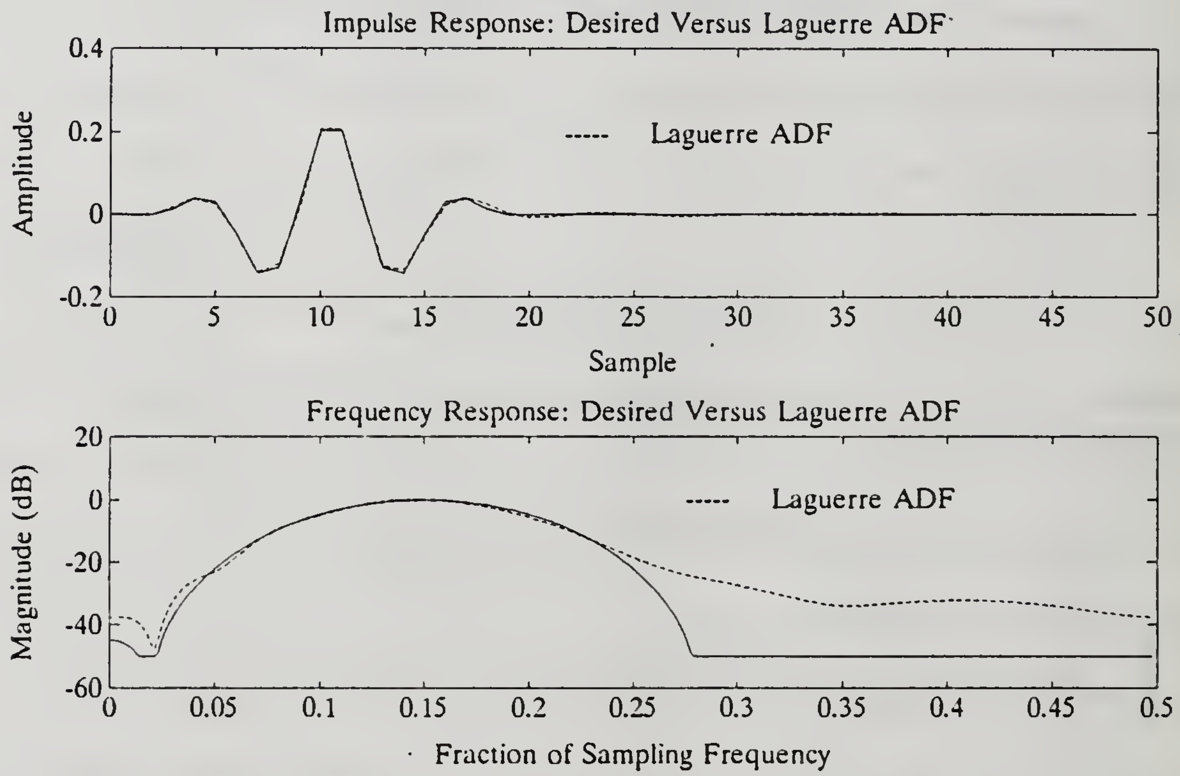

Figure 4.11: Impulse and frequency response of a 15^{th} order Laguerre ADF used to model a 7^{th} order IIR filter.

A 15^{th} order Jacobi adaptive filter with parameters $\{C,\alpha,\beta\} = \{0.0015, 4.0, 925.0\}$ was used to model the 7^{th} order IIR filter, and the plots of the model's impulse and frequency response are shown in Figure 4.13. It appears that the two additional poles in the Jacobi transfer functions do not significantly increase low order modeling performance when compared to the Legendre and Laguerre polynomials. However, there is a possibility that other parameter sets, $\{C, \alpha, \beta\}$,

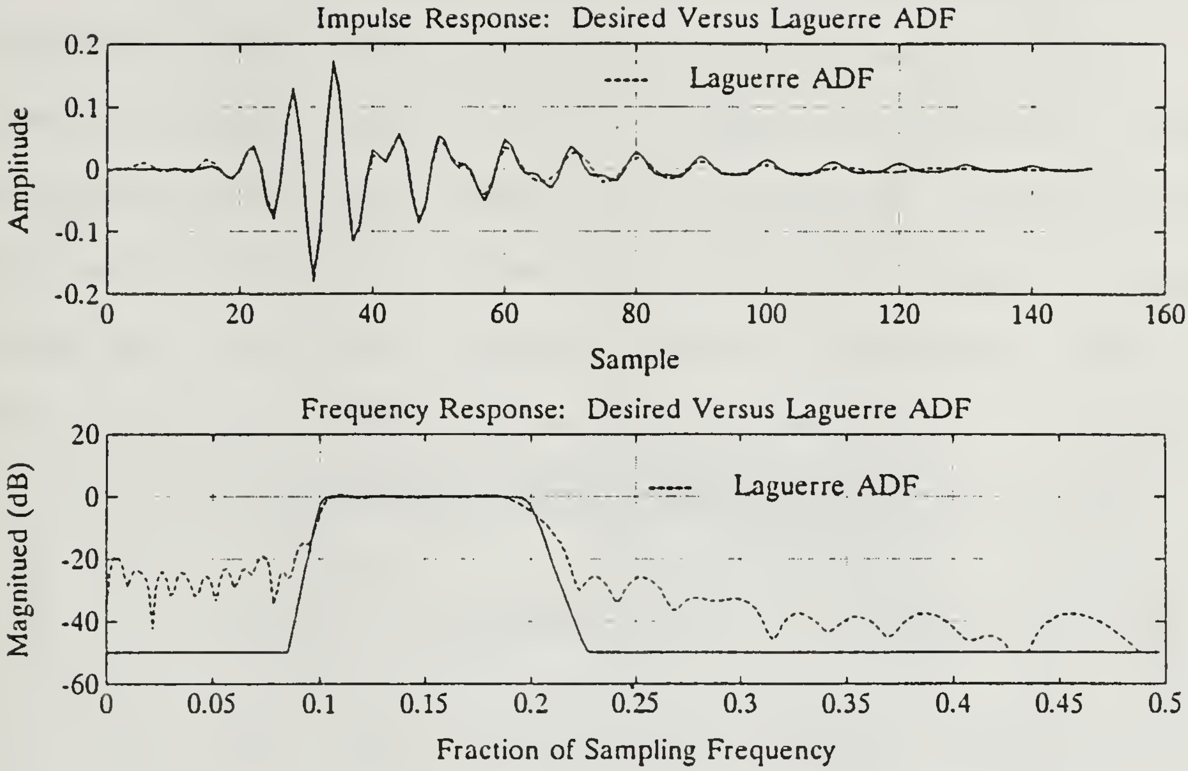

Figure 4.12: Impulse and frequency response of a 65th order Laguerre ADF used to model a 31st order IIR filter.

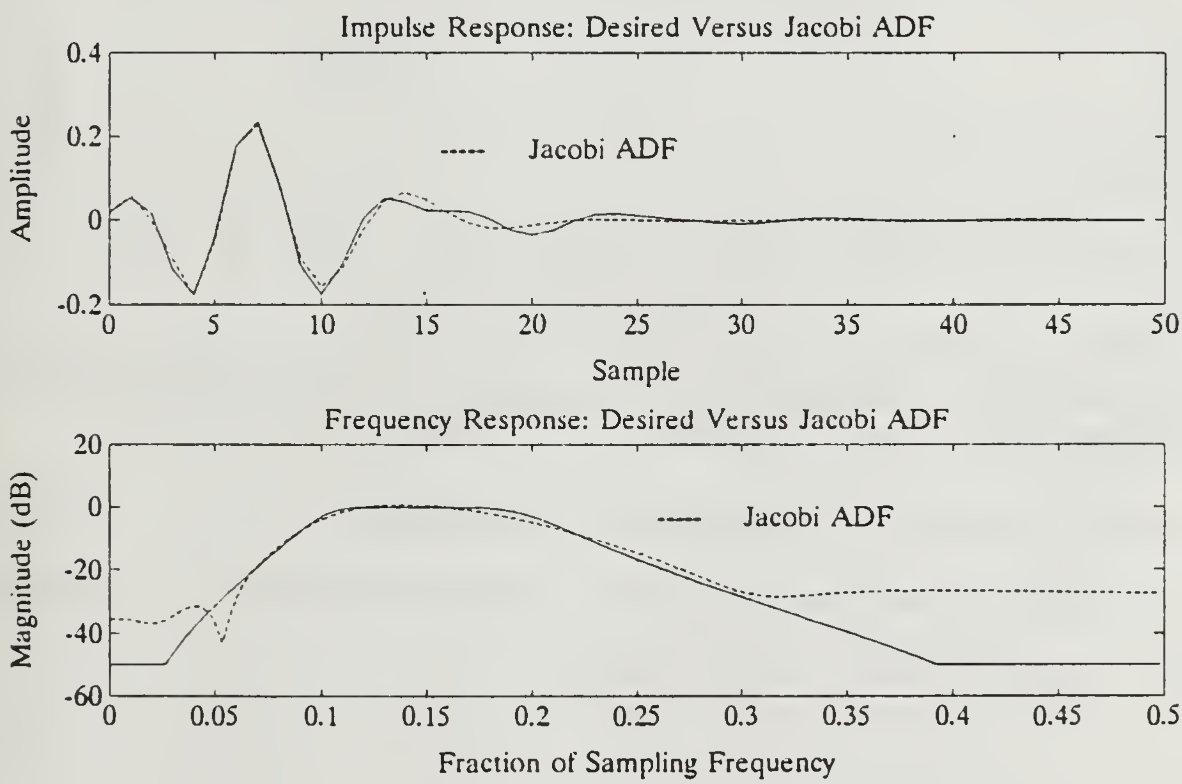

Figure 4.13: Impulse and frequency response of a 15th order Jacobi ADF used to model a 7th order IIR filter.

do exist that would produce better results.

The 15^{th} order lattice ADF simulation of the 7^{th} order IIR filter is shown in Figure 4.14. As with the FIR case, the lattice filter performs better when modeling smaller order systems. Notice the poor performance of the 65^{th} order lattice ADF when modeling the 31^{st} order IIR filter as shown in Figure 4.15. The reason the lattice filter performs worse

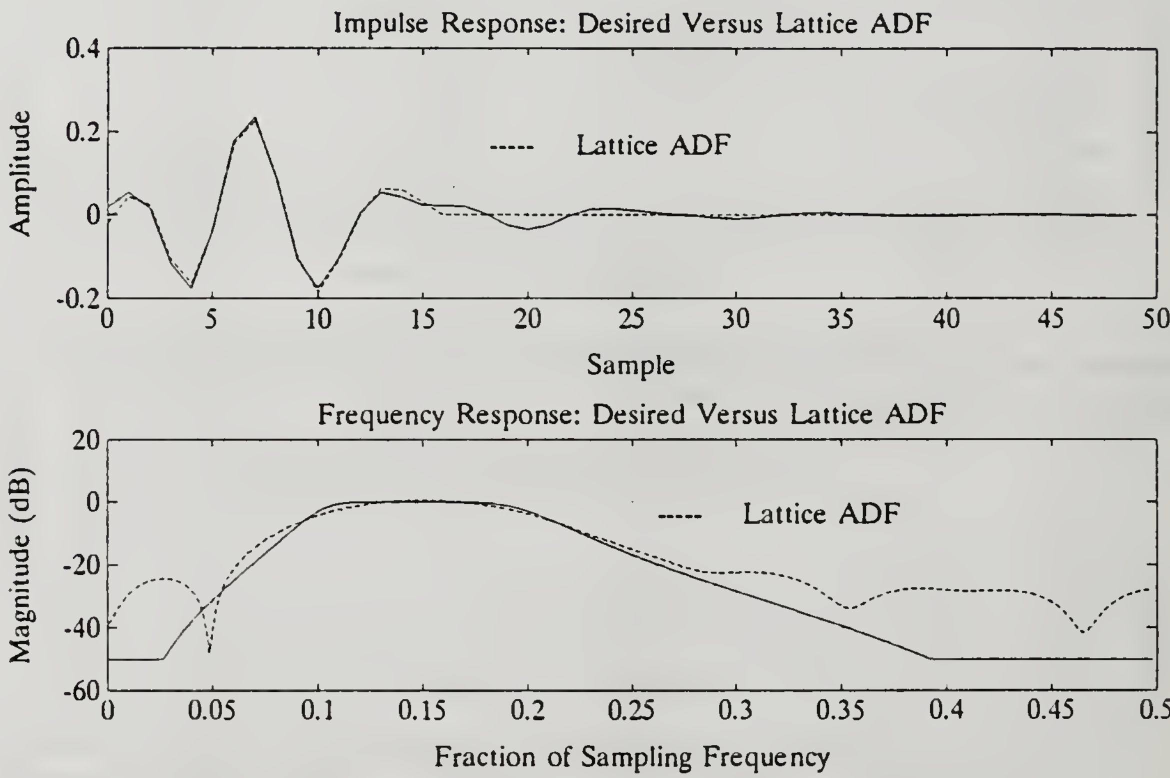

Figure 4.14: Impulse and frequency response of a 15^{th} order lattice ADF used to model a 7^{th} order IIR filter.

than the classical orthogonal polynomial filters is that it has a finite impulse response, which restricts it ability to model higher order IIR systems.

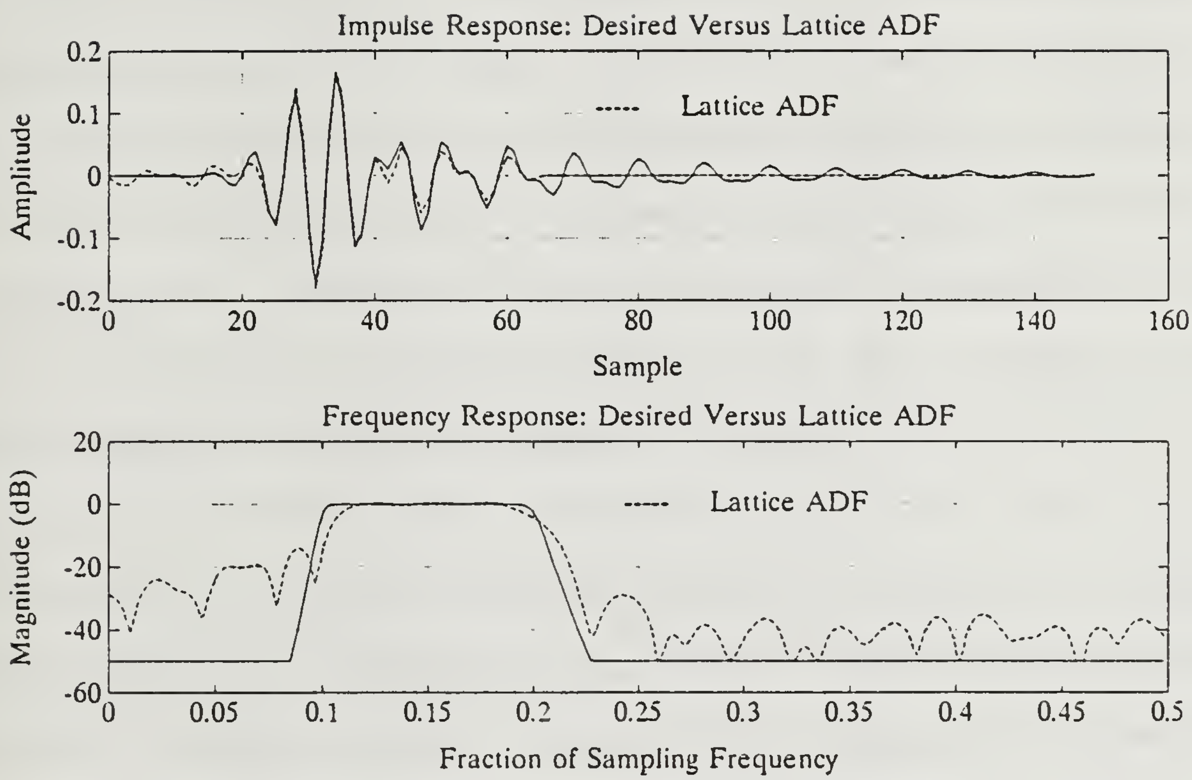

Figure 4.15: Impulse and frequency response of a 65^{th} order lattice ADF used to model a 31^{st} order IIR filter.

V. CONCLUSIONS

There are several characteristics of orthogonal filter models that make their use attractive. The classical orthogonal function filters are particularly adept at modeling systems with long impulse responses. FIR systems with large orders (say, greater than 70) can generally be modeled with 25-30% fewer coefficients using the orthogonal ADFs. Large order IIR systems can commonly be modeled with the same number of coefficients that are in the IIR system. The research in this thesis has only evaluated the performance of the orthogonal ADF models when used to identify bandpass FIR and IIR filters. It is quite possible that systems with strictly lowpass or highpass characteristics could be modeled with varying degrees of success. Although the orthogonal function filter models have shown considerable promise, there are several limitations that must be addressed; recommendations for future research are also presented.

A. LIMITATIONS OF ORTHOGONAL FILTER MODEL

Previous research has shown that the convergence rate for the LMS expansion coefficients is increased when using orthogonal functions [Ref. 1]. Even so, the length of the input sequence must be increased when the desired filter order is increased in order to allow the expansion coefficients

sufficient time to converge. Typically, the input sequence must have a length that is 10 to 15 times greater than the desired model order. This fact places limitations on the orthogonal function models, for large order models require long input sequences.

The selection of the positive real constant, C, will significantly influence the performance of the Legendre, Laguerre, and Jacobi adaptive filters. Furthermore, the Jacobi polynomials have two additional parameters, α and β, that must be chosen for any given simulation. There is currently no known algorithm for optimizing the selection of these parameters. Therefore, evaluating the performance of the orthogonal function filters is largely a matter of trial and error.

The theory for the classical orthogonal function filters was derived in the continuous time and frequency domains. The matched Z-transform technique was used to map the poles and zeros of the orthogonal filters from the s-domain to the z-domain. Unfortunately, the matched z-transform technique was found to introduce errors concerning the orthogonality of the filter output sequences. Other analog to digital transformations appear no more attractive, for their use would eliminate the ability to express the filter transfer functions in a closed form.

B. RECOMMENDATIONS FOR FURTHER RESEARCH

An algorithm to determine the selection of the positive real constant, C, is needed to reduce the trail and error nature of the classical orthogonal filter models. Further exploration of the Jacobi polynomials is bound to yield improved results, for little is known concerning the effect of altering α and β. The application of orthogonal filter models to a broad range of systems should be explored. Because the classical orthogonal polynomials are lowpass in nature, their application may be limited.

LIST OF REFERENCES

[1] H. Perez and S. Tsujii, "A System Identification Algorithm Using Orthogonal Functions," *IEEE Transactions on Signal Processing*, vol. 39, no. 3, pp. 752-754, Mar. 1991.

[2] T.W. Lee, *Statistical Theory of Communications*. New York: Wiley, 1967.

[3] P. Beckmann, *Orthogonal Polynomials for Engineers and Physicists*. Boulder, CO: The Golem Press, 1973.

[4] *Handbook of Mathematical Formulas, Tables, Graphs, Functions, Transforms*, pp. 627-63, Research and Education Association, 1980.

[5] S. Haykin, *Modern Filters*. New York: Macmillan Publishing Company, 1989.

[6] B. Widrow and S. Stearns, *Adaptive Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1985.

[7] *CRC Handbook of Mathematical Sciences*, 6th ed., CRC Press, Inc.

[8] R. Strum and D. Kirk, *First Principles of Discrete Systems and Digital Signal Processing*. Reading, MA: Addison-Wesley Publishing Company, INC, 1989.

INITIAL DISTRIBUTION LIST

		No. of Copies
1.	Defense Technical Information Center Cameron Station Alexandria, Virginia 22304-6145	2
2.	Library, Code 52 Naval Postgraduate School Monterey, CA 93943-5002	2
3.	Chairman, Code EC Department of Electrical and Computer Engineering Naval Postgraduate School Monterey, CA 93943-5000	1
4.	Professor Murali Tummala, Code EC/Tu Department of Electrical and Computer Engineering Naval Postgraduate School Monterey, CA 93943-5000	2
5.	Professor Roberto Cristi, Code EC/Cr Department of Electrical and Computer Engineering Naval Postgraduate School Monterey, CA 93943-5000	1
6.	Dr. R. Madan (Code 1114SE) Office of Naval Research 800 North Quincy Street Arlington, Virginia 22217-5000	1
7.	Mr. John Hager (Code 70E1) Naval Undersea Warfare Engineering Station Keyport, Washington 98345	1
8.	Mr. Samuel J. Frazier (Code SY84) Naval Air Warfare Center Patuxent River, Maryland 20670	1
9.	Commanding Officer Attn: LT Steven J. McPhillips, USN Naval Surface Warfare Center Dahlgren, VA 22448-5000	2

Thesis
M26948 McPhillips
c.1 Adaptive filter design using discrete orthogonal functions.